

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Lee, James
Wood, William I.

(ii) TITLE OF INVENTION: VEGF-Related Protein

(iii) NUMBER OF SEQUENCES: 12

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Genentech, Inc.
(B) STREET: 460 Point San Bruno Blvd
(C) CITY: South San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94080

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: WinPatin (Genentech)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 60/003491
(B) FILING DATE: 09/08/1995

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Lee, Wendy M.
(B) REGISTRATION NUMBER: P-40,378
(C) REFERENCE/DOCKET NUMBER: P0963R1

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 415/225-1994
(B) TELEFAX: 415/952-9881
(C) TELEX: 910/371-7168

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2031 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5 CGCGGGGTGT TCTGGTGTCC CCCGCCCCGC CTCTCCAAAA AGCTACACCG 50
 ACGCGGACCG CGGCGGCGTC CTCCCTCGCC CTCGCTTCAC CTCGCGGGCT 100
 10 CCGAATGCGG GGAGCTCGGA TGTCCGGTTT CCTGTGAGGC TTTTACCTGA 150
 CACCCGCCGC CTTTCCCCGG CACTGGCTGG GAGGGCGCCC TGCAAAGTTG 200
 GGAACGCGGA GCCCCGGACC CGCTCCCGCC GCCTCCGGCT CGCCCAGGGG 250
 15 GGGTCGCCGG GAGGAGCCCC GGGGAGAGGG ACCAGGAGGG GCCCGCGGCC 300
 TCGCAGGGGC GCCCGCGCCC CCACCCCTGC CCCC GCCAGC GGACCGGTCC 350
 20 CCCACCCCCG GTCCTTCCAC CATGCACTTG CTGGGCTTCT TCTCTGTGGC 400
 GTGTTCTCTG CTCGCCGCTG CGCTGCTCCC GGGTCCTCGC GAGGCGCCCC 450
 CCGCCGCCGC CGCCTTCGAG TCCGGACTCG ACCTCTCGGA CGCGGAGCCC 500
 25 GACGCGGGCG AGGCCACGGC TTATGCAAGC AAAGATCTGG AGGAGCAGTT 550
 ACGGTCTGTG TCCAGTGTAG ATGAACTCAT GACTGTACTC TACCCAGAAT 600
 30 ATTGGA AAT GTACAAGTGT CAGCTAAGGA AAGGAGGCTG GCAACATAAC 650
 AGAGAACAGG CCAACCTCAA CTCAAGGACA GAAGAGACTA TAAAATTTGC 700
 TGCAGCACAT TATAATACAG AGATCTTGAA AAGTATTGAT AATGAGTGGA 750
 35 GAAAGACTCA ATGCATGCCA CGGGAGGTGT GTATAGATGT GGGGAAGGAG 800
 TTTGGAGTCG CGACAAACAC CTTCTTTAAA CCTCCATGTG TGTCCGTCTA 850
 40 CAGATGTGGG GGTGCTGCA ATAGTGAGGG GCTGCAGTGC ATGAACACCA 900
 GCACGAGCTA CCTCAGCAAG ACGTTATTTG AAATTACAGT GCCTCTCTCT 950
 CAAGGCCCCA AACCAGTAAC AATCAGTTTT GCCAATCACA CTTCTGCGG 1000
 45 ATGCATGTCT AAAGTGGATG TTTACAGACA AGTTCATTCC ATTATTAGAC 1050
 GTTCCCTGCC AGCAACACTA CCACAGTGTC AGGCAGCGAA CAAGACCTGC 1100
 50 CCCACCAATT ACATGTGGAA TAATCACATC TGCAGATGCC TGGCTCAGGA 1150

AGATTTTATG TTTTCCTCGG ATGCTGGAGA TGA CTCAACA GATGGATTCC 1200
 ATGACATCTG TGGACCAAAC AAGGAGCTGG ATGAAGAGAC CTGTCAGTGT 1250
 5 GTCTGCAGAG CGGGGCTTCG GCCTGCCAGC TGTGGACCCC ACAAAGA AACT 1300
 AGACAGAAAC TCATGCCAGT GTGTCTGTAA AAACAAACTC TTCCCCAGCC 1350
 AATGTGGGGC CAACCGAGAA TTTGATGAAA ACACATGCCA GTGTGTATGT 1400
 10 AAAAGAACCT GCCCCAGAAA TCAACCCCTA AATCCTGGAA AATGTGCCTG 1450
 TGAATGTACA GAAAGTCCAC AGAAATGCTT GTTAAAAGGA AAGAAGTTCC 1500
 15 ACCACCAAAC ATGCAGCTGT TACAGACGGC CATGTACGAA CCGCCAGAAG 1550
 GCTTGTGAGC CAGGATTTTC ATATAGTGAA GAAGTGTGTC GTTGTGTCCC 1600
 TTCATATTGG AAAAGACCAC AAATGAGCTA AGATTGTACT GTTTTCCAGT 1650
 20 TCATCGATTT TCTATTATGG AAAACTGTGT TGCCACAGTA GAACTGTCTG 1700
 TGAACAGAGA GACCCTTGTG GGTCCATGCT AACAAAGACA AAAGTCTGTC 1750
 25 TTTCTGAAC CATGTGGATA ACTTTACAGA AATGGACTGG AGCTCATCTG 1800
 CAAAAGGCCT CTTGTAAAGA CTGGTTTTCT GCCAATGACC AAACAGCCAA 1850
 GATTTTCCTC TTGTGATTTT TTTAAAAGAA TGA CTATATA ATTTATTTCC 1900
 30 ACTAAAAATA TTGTTTCTGC ATTCATTTTT ATAGCAACAA CAATTGGTAA 1950
 AACTCACTGT GATCAATATT TTTATATCAT GCAAAATATG TTTAAAATAA 2000
 35 AATGAAAATT GTATTAAAA AAAAAAAAAA A 2031

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2031 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TTTTTTTTTT TTTTTTAATA CAATTTTCAT TTTATTTTAA ACATATTTTG 50
 50 CATGATATAA AAATATTGAT CACAGTGAGT TTTACCAATT GTTGTTGCTA 100

	TAAAAATGAA	TGCAGAAACA	ATATTTTITAG	TGGAAATAAA	TTATATAGTC	150
	ATTCTTTTAA	AGAAATCACA	AGAGGAAAAT	CTTGGCTGTT	TGGTCATTGG	200
5	CAGAAAACCA	GTCTTTACAA	GAGGCCTTTT	GCAGATGAGC	TCCAGTCCAT	250
	TTCTGTAAAG	TTATCCACAT	GGTTCAGGAA	AGACAGACTT	TTGTCTTTGT	300
	TAGCATGGAC	CCACAAGGGT	CTCTCTGTTC	ACAGACAGTT	CTACTGTGGC	350
10	AACACAGTTT	TCCATAATAG	AAAATCGATG	AACTGGAAAA	CAGTACAATC	400
	TTAGCTCATT	TGTGGTCTTT	TCCAATATGA	AGGGACACAA	CGACACACTT	450
15	CTTCACTATA	TGAAAATCCT	GGCTCACAAG	CCTTCTGGCG	GTTTCGTACAT	500
	GGCCGTCTGT	AACAGCTGCA	TGTTTGGTGG	TGGAACCTCT	TTCCTTTTAA	550
	CAAGCATTTT	TGTGGACTTT	CTGTACATTC	ACAGGCACAT	TTTCCAGGAT	600
20	TTAGGGGTTG	ATTTCTGGGG	CAGGTTCTTT	TACATACACA	CTGGCATGTG	650
	TTTTCATCAA	ATTCTCGGTT	GGCCCCACAT	TGGCTGGGGA	AGAGTTTGTT	700
25	TTTACAGACA	CACTGGCATG	AGTTTCTGTC	TAGTTCTTTG	TGGGGTCCAC	750
	AGCTGGCAGG	CCGAAGCCCC	GCTCTGCAGA	CACACTGACA	GGTCTCTTCA	800
	TCCAGCTCCT	TGTTTGGTCC	ACAGATGTCA	TGGAATCCAT	CTGTTGAGTC	850
30	ATCTCCAGCA	TCCGAGGAAA	ACATAAAATC	TTCCTGAGCC	AGGCATCTGC	900
	AGATGTGATT	ATTCCACATG	TAATTGGTGG	GGCAGGTCTT	GTTTCGCTGCC	950
35	TGACACTGTG	GTAGTGTGTC	TGGCAGGGAA	CGTCTAATAA	TGGAATGAAC	1000
	TTGTCTGTAA	ACATCCAGTT	TAGACATGCA	TCGGCAGGAA	GTGTGATTGG	1050
	CAAAACTGAT	TGTTACTGGT	TTGGGGCCTT	GAGAGAGAGG	CACTGTAATT	1100
40	TCAAATAACG	TCTTGCTGAG	GTAGCTCGTG	CTGGTGTTCA	TGCACTGCAG	1150
	CCCCTCACTA	TTGCAGCAAC	CCCCACATCT	GTAGACGGAC	ACACATGGAG	1200
45	GTTTAAAGAA	GGTGTTTGTC	GCGACTCCAA	ACTCCTTCCC	CACATCTATA	1250
	CACACCTCCC	GTGGCATGCA	TTGAGTCTTT	CTCCACTCAT	TATCAATACT	1300
50	TTTCAAGATC	TCTGTATTAT	AATGTGCTGC	AGCAAATTTT	ATAGTCTCTT	1350

CTGTCCTTGA GTTGAGGTTG GCCTGTTCTC TGTTATGTTG CCAGCCTCCT 1400
TTCCTTAGCT GACACTTGTA CATTTTCCAA TATTCTGGGT AGAGTACAGT 1450
5 CATGAGTTCA TCTACACTGG ACACAGACCG TAACTGCTCC TCCAGATCTT 1500
TGCTTGCATA AGCCGTGGCC TCGCCCGCGT CGGGCTCCGC GTCCGAGAGG 1550
TCGAGTCCGG ACTCGAAGGC GCGGGCGGCG GCGGGCGCCT CGCGAGGACC 1600
10 CCGGAGCAGC ACAGCGGCGA GCAGAGAACA CGCCACAGAG AAGAAGCCCA 1650
GCAAGTGCAT GGTGGAAGGA CCGGGGGTGG GGGACCGGTC CGCTGGCGGG 1700
GGCAGGGGTG GGGGCGCGGG CGCCCCTGCG AGGCCGCGGG CCCCTCCTGG 1750
15 TCCCTCTCCC CCGGGCTCCT CCCGGCGACC CCCCTGGGC GAGCCGGAGG 1800
CGGCGGGAGC GGGTCCGGGG CTCCGCGTTC CCAACTTTGC AGGGCGCCCT 1850
CCCAGCCAGT ACCGGGGAAA GCGGGCGGGT GTCAGGTAAA AGCCTCACAG 1900
GAAACCGGAC ATCCGAGCTC CCCGCATTCTG GAGCCCGCGA GGTGAAGCGA 1950
20 GGGCGAGGGA GGACGCCGCC GCGGTCCGCG TCGGTGTAGC TTTTGGAGA 2000
GGCGGGGCGG GGGACACCAG AACACCCCGC G 2031

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 amino acids
- (B) TYPE: Amino Acid
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	His	Leu	Leu	Gly	Phe	Phe	Ser	Val	Ala	Cys	Ser	Leu	Leu	Ala
1				5					10					15
Ala	Ala	Leu	Leu	Pro	Gly	Pro	Arg	Glu	Ala	Pro	Ala	Ala	Ala	Ala
				20					25					30
Ala	Phe	Glu	Ser	Gly	Leu	Asp	Leu	Ser	Asp	Ala	Glu	Pro	Asp	Ala
				35					40					45
Gly	Glu	Ala	Thr	Ala	Tyr	Ala	Ser	Lys	Asp	Leu	Glu	Glu	Gln	Leu
				50					55					60

	Arg	Ser	Val	Ser	Ser	Val	Asp	Glu	Leu	Met	Thr	Val	Leu	Tyr	Pro	
					65					70					75	
5	Glu	Tyr	Trp	Lys	Met	Tyr	Lys	Cys	Gln	Leu	Arg	Lys	Gly	Gly	Trp	
					80					85					90	
	Gln	His	Asn	Arg	Glu	Gln	Ala	Asn	Leu	Asn	Ser	Arg	Thr	Glu	Glu	
					95					100					105	
10	Thr	Ile	Lys	Phe	Ala	Ala	Ala	His	Thr	Asn	Thr	Glu	Ile	Leu	Lys	
					110					115					120	
	Ser	Ile	Asp	Asn	Glu	Trp	Arg	Lys	Thr	Gln	Cys	Met	Pro	Arg	Glu	
					125					130					135	
15	Val	Cys	Ile	Asp	Val	Gly	Lys	Glu	Phe	Gly	Val	Ala	Thr	Asn	Thr	
					140					145					150	
	Phe	Phe	Lys	Pro	Pro	Cys	Val	Ser	Val	Tyr	Arg	Cys	Gly	Gly	Cys	
					155					160					165	
	Cys	Asn	Ser	Glu	Gly	Leu	Gln	Cys	Met	Asn	Thr	Ser	Thr	Ser	Tyr	
					170					175					180	
20	Leu	Ser	Lys	Thr	Leu	Phe	Glu	Ile	Thr	Val	Pro	Leu	Ser	Gln	Gly	
					185					190					195	
	Pro	Lys	Pro	Val	Thr	Ile	Ser	Phe	Ala	Asn	His	Thr	Ser	Cys	Arg	
					200					205					210	
	Cys	Met	Ser	Lys	Leu	Asp	Val	Tyr	Arg	Gln	Val	His	Ser	Ile	Ile	
					215					220					225	
	Arg	Arg	Ser	Leu	Pro	Ala	Thr	Leu	Pro	Gln	Cys	Gln	Ala	Ala	Asn	
					230					235					240	
	Lys	Thr	Cys	Pro	Thr	Asn	Tyr	Met	Trp	Asn	Asn	His	Ile	Cys	Arg	
					245					250					255	
40	Cys	Leu	Ala	Gln	Glu	Asp	Phe	Met	Phe	Ser	Ser	Asp	Ala	Gly	Asp	
					260					265					270	
	Asp	Ser	Thr	Asp	Gly	Phe	His	Asp	Ile	Cys	Gly	Pro	Asn	Lys	Glu	
					275					280					285	
45	Leu	Asp	Glu	Glu	Thr	Cys	Gln	Cys	Val	Cys	Arg	Ala	Gly	Leu	Arg	
					290					295					300	
	Pro	Ala	Ser	Cys	Gly	Pro	His	Lys	Glu	Leu	Asp	Arg	Asn	Ser	Cys	
50					305					310					315	

Gln	Cys	Val	Cys	Lys	Asn	Lys	Leu	Phe	Pro	Ser	Gln	Cys	Gly	Ala
				320					325					330
Asn	Arg	Glu	Phe	Asp	Glu	Asn	Thr	Cys	Gln	Cys	Val	Cys	Lys	Arg
				335					340					345
Thr	Cys	Pro	Arg	Asn	Gln	Pro	Leu	Asn	Pro	Gly	Lys	Cys	Ala	Cys
				350					355					360
Glu	Cys	Thr	Glu	Ser	Pro	Gln	Lys	Cys	Leu	Leu	Lys	Gly	Lys	Lys
				365					370					375
Phe	His	His	Gln	Thr	Cys	Ser	Cys	Tyr	Arg	Arg	Pro	Cys	Thr	Asn
				380					385					390
Arg	Gln	Lys	Ala	Cys	Glu	Pro	Gly	Phe	Ser	Tyr	Ser	Glu	Glu	Val
				395					400					405
Cys	Arg	Cys	Val	Pro	Ser	Tyr	Trp	Lys	Arg	Pro	Gln	Met	Ser	
				410					415				419	

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 147 amino acids
- (B) TYPE: Amino Acid
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Asn	Phe	Leu	Leu	Ser	Trp	Val	His	Trp	Ser	Leu	Ala	Leu	Leu
1				5					10					15
Leu	Tyr	Leu	His	His	Ala	Lys	Trp	Ser	Gln	Ala	Ala	Pro	Met	Ala
				20					25					30
Glu	Gly	Gly	Gly	Gln	Asn	His	His	Glu	Val	Val	Lys	Phe	Met	Asp
				35					40					45
Val	Tyr	Gln	Arg	Ser	Tyr	Cys	His	Pro	Ile	Glu	Thr	Leu	Val	Asp
				50					55					60
Ile	Phe	Gln	Glu	Tyr	Pro	Asp	Glu	Ile	Glu	Tyr	Ile	Phe	Lys	Pro
				65					70					75
Ser	Cys	Val	Pro	Leu	Met	Arg	Cys	Gly	Gly	Cys	Cys	Asn	Asp	Glu
				80					85					90
Gly	Leu	Glu	Cys	Val	Pro	Thr	Glu	Glu	Ser	Asn	Ile	Thr	Met	Gln
				95					100					105

Ile Met Arg Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met
 110 115 120
 Ser Phe Leu Gln His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp
 125 130 135
 Arg Ala Arg Gln Glu Lys Cys Asp Lys Pro Arg Arg
 140 145 147

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 149 amino acids
 (B) TYPE: Amino Acid
 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Pro Val Met Arg Leu Phe Pro Cys Phe Leu Gln Leu Leu Ala
 1 5 10 15
 Gly Leu Ala Leu Pro Ala Val Pro Pro Gln Gln Trp Ala Leu Ser
 20 25 30
 Ala Gly Asn Gly Ser Ser Glu Val Glu Val Val Pro Phe Gln Glu
 35 40 45
 Val Trp Gly Arg Ser Tyr Cys Arg Ala Leu Glu Arg Leu Val Asp
 50 55 60
 Val Val Ser Glu Tyr Pro Ser Glu Val Glu His Met Phe Ser Pro
 65 70 75
 Ser Cys Val Ser Leu Leu Arg Cys Thr Gly Cys Cys Gly Asp Glu
 80 85 90
 Asn Leu His Cys Val Pro Val Glu Thr Ala Asn Val Thr Met Gln
 95 100 105
 Leu Leu Lys Ile Arg Ser Gly Asp Arg Pro Ser Tyr Val Glu Leu
 110 115 120
 Thr Phe Ser Gln His Val Arg Cys Glu Cys Arg Pro Leu Arg Glu
 125 130 135
 Lys Met Lys Pro Glu Arg Cys Gly Asp Ala Val Pro Arg Arg
 140 145 149

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 299 base pairs
 (B) TYPE: Nucleic Acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCGTCTACAG ATGTGGGGGT TGCTGCAATA GTGAGGGGCT GCAGTGCATG 50
 AACACCAGCA CGAGCTACCT CAGNAAGACG TTATTTGAAA TTACAGTGCC 100
 TCTCTCTCAA GGCCCCAAAC CAGTAACAAT CAGTTTTGCC AATCACACTT 150
 CCTGCCGATG CATGTCTAAA CTGGATGTTT ACAGACAAGT TCATTCCATT 200
 ATTAGACGTT CCCTGCCAGC AACACTACCA CAGTGTCAGG CAGCGAACAA 250
 GACCTGCCCC ACCAATTACA TGTGGAATAA TCACATCTGC AGATGCCTG 299

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 50 base pairs
 (B) TYPE: Nucleic Acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTGGTGTTCA TGCACTGCAG CCCCTCACTA TTGCAGCAAC CCCCACATCT 50

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 50 base pairs
 (B) TYPE: Nucleic Acid
 (C) STRANDEDNESS: Single
 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCATCTGCAG ATGTGATTAT TCCACATGTA ATTGGTGGGG CAGGTCTTGT 50

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Tyr Ser Met Thr Pro Pro Thr Leu
1 5 8

10 (2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: Amino Acid
15 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ser Leu Arg Arg Arg Gln Gln Gln Asp
1 5 9

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 amino acids
(B) TYPE: Amino Acid
25 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Lys Tyr Ala Leu Ala Asp Ala Ser Leu Lys Met Ala Asp Pro Asn
1 5 10 15
Arg Phe Arg Gly Lys Asp Leu Pro Val Leu Asp Gln Leu Leu Glu
20 25 30
Gly Gly Ala Ala His Tyr Ala Leu Leu Pro
35 40

40 (2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: Amino Acid
45 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly Pro Arg Glu Ala Pro Ala Ala Ala Ala Phe Glu
1 5 10 13